

# Measuring BR ( $h \rightarrow \tau^+\tau^-$ ) at the ILC: a full simulation study

Shin-ichi Kawada<sup>1,†</sup>, Keisuke Fujii<sup>2</sup>, Taikan Suehara<sup>3</sup>,  
Tohru Takahashi<sup>1</sup>, Tomohiko Tanabe<sup>4</sup>, Harumichi Yokoyama<sup>5</sup>

March 27, 2015

- 1: Graduate School of Advanced Sciences of Matter (AdSM), Hiroshima University, 1-3-1, Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8530, Japan  
 2: High Energy Accelerator Research Organization (KEK), 1-1, Oho, Tsukuba, Ibaraki, 305-0801, Japan  
 3: Graduate School of Science, Kyushu University, 6-10-1, Hakozaki, Higashi-ku, Fukuoka, 812-8581, Japan  
 4: International Center for Elementary Particle Physics (ICEPP), The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan  
 5: Graduate School of Science, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan

† : s-kawada@huhep.org

## Abstract

<sup>1</sup> We evaluate the expected measurement accuracy of the branching ratio of the Standard Model Higgs boson decaying into tau lepton pairs at the ILC with a full simulation of the ILD detector concept. We assume a Higgs mass of 125 GeV, a branching ratio of  $\text{BR}(h \rightarrow \tau^+\tau^-) = 6.32\%$ , a beam polarization of electron (positron) of  $-0.8(+0.3)$ , and an integrated luminosity of  $250 \text{ fb}^{-1}$ . The Higgs-strahlung process  $e^+e^- \rightarrow Zh$  with  $Z \rightarrow q\bar{q}$  is analyzed. We estimate the measurement accuracy of the branching ratio  $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR})$  to be 3.4% with using a multivariate analysis technique.

## 1 Introduction

The Higgs boson has been found at the LHC by the ATLAS experiment [1] and the CMS experiment [2]. After that, understanding the properties of a Higgs boson through the precise measurement is important for particle physics. A branching ratio (BR) is one of the most important properties of a Higgs boson. In the Standard Model, the Yukawa coupling of matter fermions with a Higgs boson is completely proportional to the fermion mass. However, the Yukawa coupling will deviate from the prediction of the Standard Model if there are new physics. The pattern of deviation depends on the new physics model. Thus, understanding a Higgs boson is very important from the viewpoint of new physics.

Besides, the size of the deviation is expected to be small if the scale of new physics is high. Specifically, the allowed deviation can be at the few-percent level even if no additional new particles are to be found at the LHC [3]. Since the branching ratio measurement is used as an input in the extraction of the Yukawa coupling, a precise determination of the branching ratio is essential to probe new physics.

In this study, we focus on the branching ratio of the Higgs boson decaying into tau lepton pairs. The mass of the tau lepton is known to a very good precision unlike quarks, which typically suffer from the theoretical uncertainties arising from QCD. Also, the deviation in the lepton Yukawa coupling could well differ from the quark Yukawa coupling, such as in the lepton-specific Two-Higgs Doublet Model. Thus, the tau Yukawa coupling is an ideal probe for new physics.

In this study, we estimate the measurement accuracy  $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR})$  of the  $h \rightarrow \tau^+\tau^-$  branching ratio at the center-of-mass energy of 250 GeV and 500 GeV at the ILC with the ILD full detector simulation. In this proceedings, we will discuss the 250 GeV case, based on our talk at the LCWS14 [4].

## 2 Signal and Background

The Feynman diagrams of the Higgs boson production processes are shown in Figure 1. The Higgs-strahlung process ( $e^+e^- \rightarrow Zh$ ) is the most dominant production processes at the center-of-mass energy ( $\sqrt{s}$ ) of 250 GeV. We analyze this process with the  $Z$  boson decaying into quark pair ( $Z \rightarrow q\bar{q}$ ), which is expected to be the most sensitive channel because of the high statistics. The cross section of the Higgs-strahlung process at  $\sqrt{s} = 250 \text{ GeV}$  is about 300 fb [5].

<sup>1</sup>Talk presented at the International Workshop on Future Linear Colliders (LCWS14), Belgrade, Serbia, 6-10 October 2014.

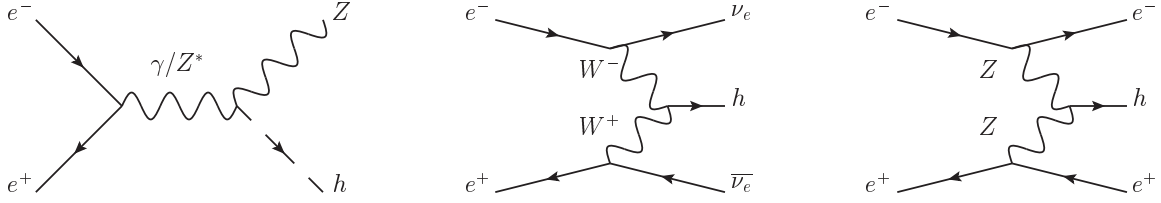


Figure 1: The Feynman diagrams of the Higgs boson production processes. (left): Higgs-strahlung process, (middle):  $WW$ -fusion process, (right):  $ZZ$ -fusion process.

The diagrams for the main backgrounds which have the same(similar) final state as the signal are shown in Figure 2. The  $e^+e^- \rightarrow ZZ \rightarrow q\bar{q}\tau^+\tau^-$  process shown in left of Figure 2 is an irreducible background to the signal. Besides, the semi-leptonic decay of di-boson processes such as  $e^+e^- \rightarrow W^+W^- \rightarrow q\bar{q}'\tau\nu$  shown in Figure 2 right will also to be the source of backgrounds due to the similar final states to the signal.

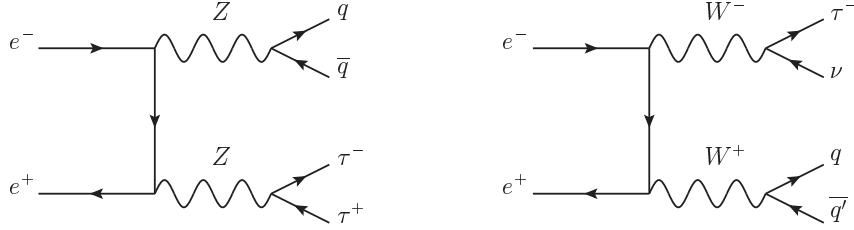


Figure 2: The example of Feynman diagrams of the main backgrounds.

### 3 Simulation Conditions

We assume a Higgs mass of  $M_h = 125$  GeV, a branching ratio of  $\text{BR}(h \rightarrow \tau^+\tau^-) = 6.32\%$  [6], an integrated luminosity of  $250 \text{ fb}^{-1}$ , and beam polarizations of  $P(e^-, e^+) = (-0.8, +0.3)$ .

We use the Monte-Carlo samples which are prepared for the studies presented in the ILC Technical Design Report [5,7–9] (so called TDR sample). Besides, we use the event generator [10] with using GRACE [11,12] and generate Monte-Carlo samples of the processes of  $e^+e^- \rightarrow f\bar{f}h$  with  $h \rightarrow \tau^+\tau^-$  where  $f$  denotes a fermion, because the spin correlation of  $h \rightarrow \tau^+\tau^-$  was not treated properly in TDR samples. Therefore, we do not use the events which contain  $h \rightarrow \tau^+\tau^-$  processes in TDR samples to avoid double counting. The beam energy spectrum includes the effects due to beamstrahlung and the initial state radiation. The beam-induced backgrounds from  $\gamma\gamma$  interactions which give rise to hadrons are included in all signal and background processes. The background processes from  $e^+e^-$  interactions are categorized according to the number of final-state fermions: two fermions (2f), three fermions (1f\_3f), and four fermions (4f). We also include  $\gamma\gamma \rightarrow 2f$  and  $\gamma\gamma \rightarrow 4f$  processes for this study. The detector response is simulated using full simulation based on Geant4 [13]. We perform the detector simulation with Mokka [15], a Geant4-based full simulation, with the ILD detector model of ILD\_o1\_v05. TAUOLA [16] is used for the tau decay simulation. The ILD detector model consists of a vertex detector, a time projection chamber, an electromagnetic calorimeter, a hadronic calorimeter, a return yoke, muon systems, and forward components.

### 4 Event Reconstruction

Even in the signal processes, some hadrons generated via the beam-induced  $\gamma\gamma$  interactions are included. But at the  $\sqrt{s} = 250$  GeV, the numbers of hadrons are estimated to be small [17]. Therefore, we do not apply any special treatment to eliminate these hadrons.

The event reconstruction is consisted of three steps; (1) tau reconstruction, (2) collinear approximation, and (3)  $Z$  boson reconstruction.

First, we apply tau finder. Our tau finder searches for charged track which have highest energy among the remaining particles, and combines the neighboring particles which satisfy  $\cos\theta_{\text{cone}} > 0.99$ , with the combined mass less than 2 GeV, where  $\theta_{\text{cone}}$  is the cone angle with respect to the highest energy charged track. The combined object is regarded as a tau candidate. Then, we apply the selection cuts to tau candidates as following; discard the candidates which categorized into 3-prong with neutral particles,  $E_{\text{tau candidate}} > 3$  GeV where  $E_{\text{cand}}$  is the energy of tau candidate, and satisfy  $E_{\text{cone}} < (0.1 \times E_{\text{cand}})$  with  $\cos\theta'_{\text{cone}} = 0.95$  where  $E_{\text{cone}}$  is the cone energy of a tau candidate with the cone angle of  $\theta'_{\text{cone}}$ , and  $\theta'_{\text{cone}}$  is the cone angle with respect to a tau candidate, respectively. These selections are tuned to minimize the misidentification of fragments of quark jets as tau decays. After the selection, we apply the charge recovery process

to obtain better efficiency. The charged particles in a tau candidate which have the energy less than 2 GeV are detached one by one, the one with the smallest energy first, until satisfying the following conditions; the charge of a tau candidate is exactly equal to  $\pm 1$ , and the number of track(s) in a tau candidate is exactly equal to 1 or 3. The tau candidate after the detaching is rejected if it does not satisfy the above conditions, and remaining candidate is regarded as a tau jet. We repeat the above processes until there are no charged particles which have the energy greater than 2 GeV.

After finishing the tau reconstruction, we apply the collinear approximation [18] to reconstruct the invariant mass of tau pair system. In this approximation, we assume that the visible decay products of the tau lepton and the neutrino(s) from the tau decay is collinear, and the contribution of the missing transverse momentum comes only from the neutrino(s) from tau decay.

After the approximation, we apply the Durham jet clustering [19] with two jets for the remaining objects to reconstruct  $Z$  boson.

## 5 Event Selection

Before optimizing the cuts, we apply preselection cuts as following;

- the number of reconstructed  $\tau^+$  and  $\tau^-$  are exactly equal to one,
- the number of reconstructed jet is exactly equal to two,
- the number of track in an event is greater or equal to nine,
- $M_{\text{col}} > 0$  and  $E_{\text{col}} > 0$  where  $M_{\text{col}}(E_{\text{col}})$  is the invariant mass (energy) of tau pair with collinear approximation in the unit of GeV.

In addition, we apply the following cuts to suppress the trivial backgrounds;

- $105 < E_{\text{vis}} < 215$  and  $M_{\text{vis}} > 95$ , where  $E_{\text{vis}}(M_{\text{vis}})$  is the visible energy (mass) in the unit of GeV,
- $P_t > 40$ , where  $P_t$  is the sum of the magnitude of the transverse momentum of each particle in the unit of GeV,
- the thrust in an event should be less than 0.97,
- $60 < E_{q\bar{q}} < 175$  and  $35 < M_{q\bar{q}} < 160$ , where  $E_{q\bar{q}}(M_{q\bar{q}})$  is the energy (invariant mass) of reconstructed two jets in the unit of GeV,
- $\cos \theta_{q\bar{q}} < 0.5$ , where  $\theta_{q\bar{q}}$  is the angle between two jets,
- $E_{\tau^+\tau^-} < 140$  and  $5 < M_{\tau^+\tau^-} < 125$ , where  $E_{\tau^+\tau^-}(M_{\tau^+\tau^-})$  is the energy (invariant mass) of reconstructed two taus without collinear approximation in the unit of GeV,
- $\cos \theta_{\tau^+\tau^-} < -0.1$ , where  $\theta_{\tau^+\tau^-}$  is the angle between two taus,
- $30 < E_{\text{col}} < 270$  and  $15 < M_{\text{col}} < 240$ ,
- $65 < M_{\text{recoil}} < 185$ , where  $M_{\text{recoil}}$  is the recoil mass against  $Z$  boson.

We only use the events passed the preselection and cuts above for the optimization of event selection. We perform the cut-based analysis and multivariate analysis both. The optimization is performed to maximize the signal statistical significance of  $S/\sqrt{S+B}$ , where  $S(B)$  is the number of signal(background).

### 5.1 Cut-based Analysis

We apply the following cuts sequentially to extract maximum signal significance;

- $E_{\text{vis}} < 240$  GeV,
- $|\cos \theta_{\text{miss}}| < 0.98$ , where  $\theta_{\text{miss}}$  is the angle between missing momentum in an event and beam axis,
- $E_{q\bar{q}} < 125$  GeV,  $M_{q\bar{q}} > 80$  GeV,
- $E_{\tau^+\tau^-} < 130$  GeV,  $M_{\tau^+\tau^-} < 115$  GeV,  $\cos \theta_{\tau^+\tau^-} < 0.54$ ,
- $E_{\text{col}} < 210$  GeV,  $M_{\text{col}} > 100$  GeV,
- $\log_{10} |d_0 \text{sig}(\tau^+)| + \log_{10} |d_0 \text{sig}(\tau^-)| > -0.2$ , where  $d_0 \text{sig}(\tau^+)$  is the impact parameter  $d_0$  ( $xy$ -plane) divided by the error of  $d_0$  for  $\tau^+$ ,
- $\log_{10} |z_0 \text{sig}(\tau^+)| + \log_{10} |z_0 \text{sig}(\tau^-)| > -0.2$ , where  $z_0 \text{sig}(\tau^+)$  is the impact parameter  $z_0$  ( $rz$ -plane) divided by the error of  $z_0$  for  $\tau^+$ ,

- $M_{\text{recoil}} > 115 \text{ GeV}$ .

Figure 3 shows the  $M_{\text{col}}$  distribution in the sequential cuts. After all cuts above, the signal events of 1002 and the background events of 535.4 are remained. The statistical significance is calculated to be  $S/\sqrt{S+B} = 25.6\sigma$ . This result corresponds to the precision of  $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR}) = 3.9\%$ .

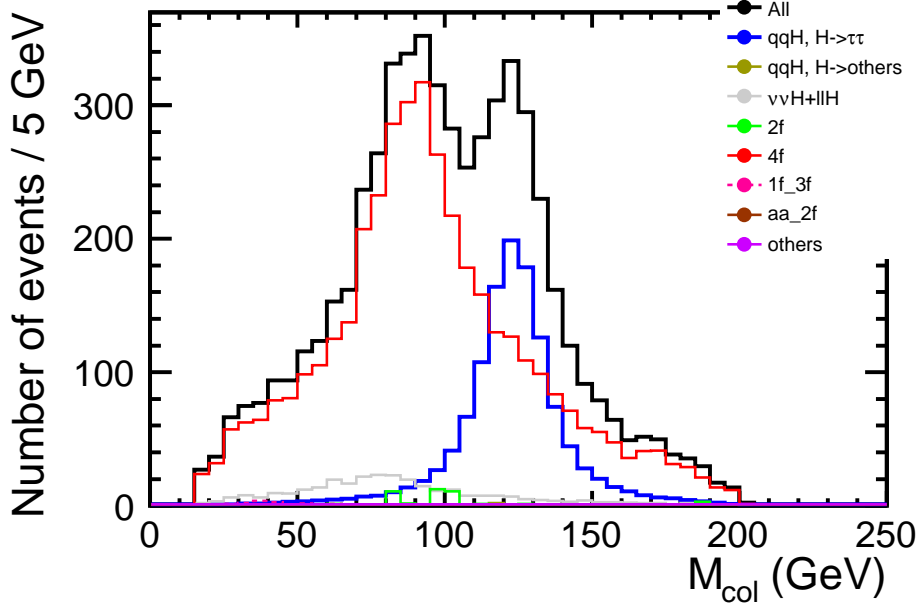


Figure 3: The distribution of  $M_{\text{col}}$  in the sequential cuts. Black histogram shows the summing up of all processes, blue shows the signal process, and red shows the all four-fermion background processes, respectively.

## 5.2 Multivariate Analysis

We use the TMVA package in ROOT [20] and use the Gradient Boosted Decision Tree (BDTG) technique for the analysis tool. We use the following 17 variables as the inputs;

- three variables from the overall event:  $E_{\text{vis}}$ ,  $P_t(\text{all})$ , and  $\cos \theta_{\text{miss}}$ , where  $P_t(\text{all})$  is the magnitude of the transverse momentum calculated from the momentum vector of all visible particles in an event,
- three variables from reconstructed jets and  $Z$  boson:  $M_{q\bar{q}}$ ,  $\cos \theta_{q\bar{q}}$ , and  $\cos \theta_Z$  where  $\theta_Z$  is the angle between the momentum of reconstructed  $Z$  boson and beam axis,
- seven variables from reconstructed taus:  $M_{\tau^+\tau^-}$ ,  $E_{\tau^+\tau^-}$ ,  $\cos \theta_{\tau^+\tau^-}$ ,  $\cos \theta_{\text{acop}}$ ,  $\cos \theta_H$ ,  $\log_{10} |d_0 \text{sig}(\tau^+)| + \log_{10} |d_0 \text{sig}(\tau^-)|$ ,  $\log_{10} |z_0 \text{sig}(\tau^+)| + \log_{10} |z_0 \text{sig}(\tau^-)|$ , where  $\theta_{\text{acop}}$  is the acoplanarity angle between two taus,  $\theta_H$  is the angle between the momentum of reconstructed Higgs boson without collinear approximation and beam axis,
- three variables from collinear approximation:  $M_{\text{col}}$ ,  $E_{\text{col}}$ ,  $\cos \theta_{H\text{col}}$  where  $\theta_{H\text{col}}$  is the angle between reconstructed Higgs momentum and beam axis with collinear approximation,
- one from recoil mass:  $M_{\text{recoil}}$ .

Figure 4 shows the response of the multivariate classifier. The number of events surviving the event selection is 1205 for the signal and 521.4 for the background. The signal significance is calculated to be  $S/\sqrt{S+B} = 29.0\sigma$ , which implies a precision of  $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR}) = 3.4\%$ . The result of the multivariate analysis improved by  $\sim 13\%$  than the cut-based analysis.

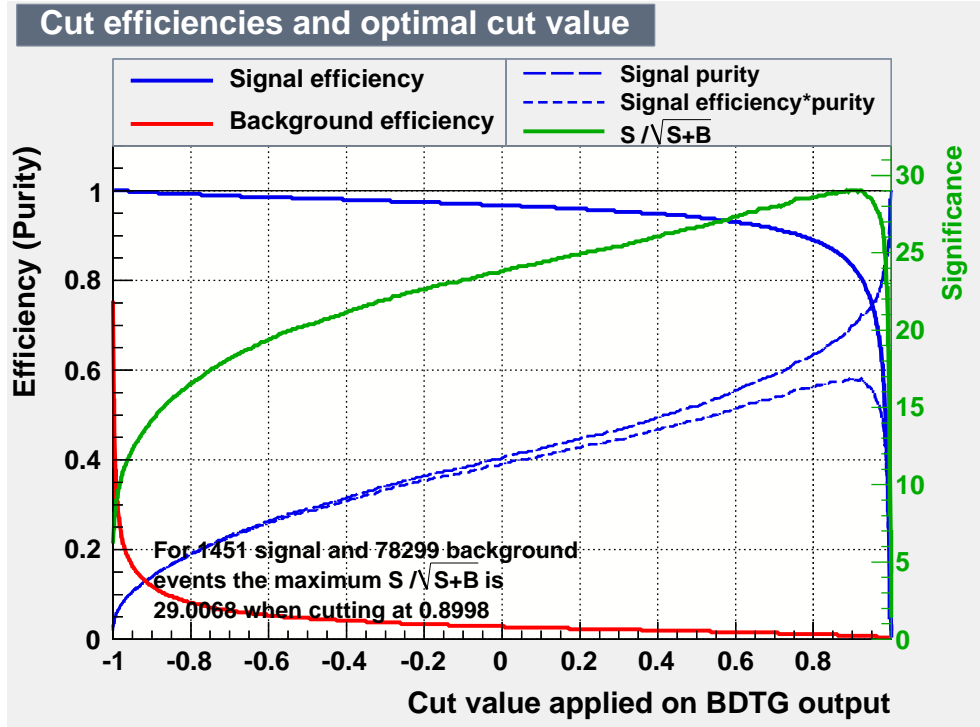


Figure 4: The multivariate classifier response for  $q\bar{q}h$  mode. The green plot shows the signal significance.

## 6 Comparison with Previous Results

We previously performed the analysis with fully simulated samples at  $\sqrt{s} = 250$  GeV, and the results are written in Ref [21]. In Ref [21], the analysis had been performed with the condition of Higgs mass of 120 GeV, and analyzed three signal modes:  $q\bar{q}h$ ,  $e^+e^-h$ , and  $\mu^+\mu^-h$ . The combined result with the extrapolation to the Higgs mass of 125 GeV was  $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR}) = 4.2\%$ . The results described in Section 5 are much better than previous study, even the cross section and the branching ratio of  $h \rightarrow \tau^+\tau^-$  [6] are worse than the case of Higgs mass of 120 GeV and only using  $q\bar{q}h$  mode. These differences mainly come from the mass difference between  $Z$  boson and Higgs boson, and the difference of analysis technique. Now we can apply tighter cuts in the recoil mass against  $Z$  boson than previously, because the peak of recoil mass shifted from 120 GeV to 125 GeV. This difference introduces good separation power between signal and background. Besides, we only had been performed cut-based analysis in previous. The cut-based analysis is very simple and easy to understand, but difficult to get better precision. The multivariate analysis gives us much better results than the cut-based, as described on Section 5.2.

## 7 Summary

We evaluated the expected measurement accuracy of the branching ratio  $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR})$  of the  $h \rightarrow \tau^+\tau^-$  mode at  $\sqrt{s} = 250$  GeV at the ILC with a full simulation of the ILD detector model, assuming  $M_h = 125$  GeV,  $\text{BR}(h \rightarrow \tau^+\tau^-) = 6.32\%$ ,  $\int L dt = 250 \text{ fb}^{-1}$ , and beam polarizations  $P(e^-, e^+) = (-0.8, +0.3)$ . We analyzed the  $q\bar{q}h$  final state using a cut-based approach and a multivariate approach. As a result, we expected  $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR}) = 3.4\%$  with using Gradient Boosted Decision Tree technique.

## References

- [1] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716** (2012) 1 - 29
- [2] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716** (2012) 30 - 61
- [3] R. S. Gupta, H. Rzehak, J. D. Wells, Phys. Rev. D **86**, 095001 (2012)
- [4] Shin-ichi Kawada, Tomohiko Tanabe, Taikan Suehara, Keisuke Fujii, Tohru Takahashi, Harumichi Yokoyama, "Measuring  $\text{BR}(h \rightarrow \tau^+\tau^-)$  at the ILC: a full simulation study", talk at the LCWS14 (2014)
- [5] The International Linear Collider Technical Design Report Volume 2: Physics (2013)

- [6] S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka, *et al.* [LHC Higgs Cross Section Working Group], arXiv:1201.3084v1 [hep-ph] (2012)
- [7] The International Linear Collider Technical Design Report Volume 1: Executive Summary (2013)
- [8] The International Linear Collider Technical Design Report Volume 3: Accelerator (2013)
- [9] The International Linear Collider Technical Design Report Volume 4: Detectors (2013)
- [10] Harumichi Yokoyama's master thesis, (2014)
- [11] F. Yuasa, J. Fujimoto, T. Ishikawa, M. Jimbo, T. Kaneko, K. Kato, S. Kawabata, T. Kon, Y. Kurihara, M. Kuroda, N. Nakazawa, Y. Shimizu and H. Tanaka, Prog. Theor. Phys. Suppl. **138** 182000 arXiv:hep-ph/0007053
- [12] "Minami-Tateya web page", <http://www-sc.kek.jp>
- [13] S. Agostinelli *et al.* [GEANT4 Collaboration], Nucl. Instrum. Meth. A **506**, 250 - 303 (2003).
- [14] <http://berggren.home.cern.ch/berggren/sgv.html>
- [15] P. Mora de Fretias, H. Videau, LC-TOOL-2003-010 (2003)
- [16] S. Jadach, J. H. Kühn, Z. Was, Comput. Phys. Commun. **64** (1991) 275 - 299
- [17] Pisin Chen, Timothy L. Barklow, and Michael E. Peskin, Phys. Rev. D **49** (1994) 3209 - 3227
- [18] R. K. Ellis, I. Hinchliffe, M. Soldate, J. J. van der bij, Nucl. Phys. B **297** (1988) 221 - 243
- [19] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, B. R. Webber, Phys. Lett. B **269** (1991) 432 - 438
- [20] <http://root.cern.ch/>
- [21] Shin-ichi Kawada, Keisuke Fujii, Taikan Suehara, Tohru Takahashi, Tomohiko Tanabe, LC-Note, LC-REP-2013-001